LEARNING BY LIKING -
A MERE EXPOSURE VERSION OF THE AGL PARADIGM

Master’s thesis in Cognitive Science

Author: Åsa Elwér, asael092@student.liu.se
Supervisors: Martin Ingyar, Karl-Magnus Petersson and Christian Forkstam
ISRN: LIU-KOGVET-D—04/09—SE
Date: 2004-06-21
Learning by liking – a Mere Exposure version of the AGL paradigm
Learning by liking – a Mere Exposure version of the AGL paradigm

Abstract

The artificial grammar learning (AGL) paradigm has been intensively researched since the 60-s. In general, these investigations attempt to study the implicit acquisition of structural regularities. Among other things, it has been suggested that the AGL paradigm can serve as a model for the process of acquiring a natural language. Thus it can serve as a well-controlled laboratory task that might be used to understand certain aspects of the process of language acquisition. For example the AGL paradigm has been used in an attempt to isolate the acquisition of syntactic aspects of language. Several experimental studies show that the participants acquire knowledge of the underlying rule system since they are able to differentiate grammatical strings from non-grammatical ones. It has been argued that the traditionally conducted AGL paradigm with grammaticality instructions might make the task explicit, at least during the test phase. In order to imitate the language learning process as close as possible, to rule out the possibility of an explicit component during the testing phase (i.e., keeping the retrieval process implicit) and to rule out explicit rule conformity or rule following, we modified the classical AGL paradigm. In a behavioural study we combined the AGL paradigm with an altered mere exposure paradigm in an attempt to better model aspects of language acquisition. We were able to show that subjects, classifying under mere exposure instructions, categorize grammatical and non-grammatical strings just as well as those solving the classification task with the grammaticality instructions. This indicates that the mere exposure version might serve as a more appropriate model for language acquisition.
# Table of contents

ABSTRACT ............................................................................................................... 3  
TABLE OF CONTENTS .................................................................................... 4  
1. INTRODUCTION ........................................................................................... 6  
2. THEORY .......................................................................................................... 9  
   MEMORY SYSTEMS ............................................................................................. 9  
   ARTIFICIAL GRAMMAR LEARNING .............................................................. 12  
   FORMAL LANGUAGES ....................................................................................... 13  
   REPRESENTATION SYSTEMS .............................................................................. 15  
   EXPLICIT OR IMPLICIT PROCESSING .......................................................... 16  
   THE MERE EXPOSURE PARADIGM ............................................................... 18  
   ASSOCIATIVE CHUNK STRENGTH ...................................................................... 19  
   CONTROL GROUPS IN AGL EXPERIMENTS ...................................................... 20  
   AGL AND LANGUAGE ACQUISITION ............................................................. 21  
   THE EXPERIMENT ............................................................................................. 22  
3. MATERIALS AND METHODS ........................................................................ 23  
   SUBJECTS .......................................................................................................... 23  
   STIMULUS MATERIAL ........................................................................................ 23  
   EXPERIMENTAL GROUPS ................................................................................... 24  
   INSTRUCTIONS ................................................................................................... 26  
   EXPERIMENTAL PROCEDURE ......................................................................... 27  
   DATA ANALYSIS ............................................................................................... 27  
4. RESULTS ....................................................................................................... 29  
   QUALITATIVE DATA .......................................................................................... 34  
      Acquisition Phase ......................................................................................... 34  
      Classification Phase ..................................................................................... 34  
5. DISCUSSION ................................................................................................. 36  
   INITIAL BIAS ..................................................................................................... 36  
   LEARNING .......................................................................................................... 37  
   THE MERE EXPOSURE PARADIGM ............................................................... 38  
   REPRESENTATION SYSTEMS .............................................................................. 39  
6. CONCLUSIONS ............................................................................................ 41  
REFERENCES .................................................................................................. 42  
APPENDIX A: GRAMMATICAL ACQUISITION STRINGS ................... 44  
APPENDIX B: RANDOMIZED ACQUISITION STRINGS ....................... 45
1. Introduction

Since the 1950’s many attempts have been made to create a model for investigating the process of language learning. One approach devised to study this issue is based on the use of stimulus material generated from an artificial grammar in an implicit learning paradigm. In a seminal paper by Reber (Reber 1967), he described the paradigm Artificial Grammar Learning (AGL). A typical AGL experiment includes an acquisition phase and a classification phase. During the acquisition phase, participants are engaged in a short-term memory task using an acquisition sample of symbol sequences generated from an artificial grammar, commonly a finite state machine. The finite state machine implements a set of rules, that is, the rules of the grammar. Subsequent to the acquisition phase the subjects are informed that the items (i.e., symbol sequences) are generated according to a complex system of rules and they were asked to classify new items, not previously encountered, as grammatical or non-grammatical, guided by their immediate intuitive impression (‘gut feeling’). Typically, subjects perform reliably above chance on this task, in other words, the subjects can in the decoding session be shown to have acquired some form of knowledge or aspects of the underlying rules that generate structural regularities in the observed strings. As subjects were unable to motivate their decisions it was presumed that it was based on implicit retrieval. The AGL paradigm has been proposed as a relevant laboratory model for language acquisition in infants (Gomez and Gerken 1999) and second language learning in adults (Friederici et al., 2002).

In the AGL paradigm we try to isolate the syntactic aspects of language and use a formal language defined by a finite automaton as a model for natural language. One component in the definition of a formal language is its finite lexicon (alphabet) $V$ of terminal symbols, $V = \{t_1, t_2, ..., t_N\}$. The set of all possible finite symbol strings that can be generated from the alphabet $V$ is given by Kleene-star operator $V^* = \{\emptyset, t_1, t_2, ..., t_N, t_1t_1, t_1t_2, t_1t_3, ..., tk_1tk_2...tk_m, ...\}$. A formal language $L$ over $V$ is then defined as a subset of $V^*$, $L \subseteq V^*$; a symbol string $s = tk_1tk_2...tk_m$ is well-formed or grammatical if and only if $s \in L$. (Lewis and Papadimitriou 1981) This way of introducing formal languages amounts to an extensional definition, an E-language, where the language is identified with its string set. This is adequate for formal investigations but may, perhaps, be of limited interest from a cognitive point of view. In the context of natural language grammars it is questionable whether an extensional definition is at all meaningful. Instead, a more interesting approach, takes as its point of departure an intentional definition of language. This amounts to the specification of a generating mechanism, including principles of combinations and additional non-terminal symbols, capable of generating all grammatical (well-formed) strings and only those in a given language. The generating mechanism serves as
an intentional definition of the language, an I-language, and a string \( s \) is grammatical (\( s \in L \)) if and only if the formal mechanism (or machine) can generate it. (Chomsky 2000) Here, it should be noted that the term ‘language’ in formal language, do not entail anything specific beyond what is outlined above, and that a formal (or artificial) grammar represents a formal specification of a mechanism that generate (or recognize) certain types of structural regularities. (Petersson et al in press)

Successful artificial grammar learning has been shown in many experiments. Subjects seem to learn to categorize the strings according to the underlying grammar. Since we use the AGL experiment as a model for language acquisition we want it to imitate the language learning process as close as possible. The language learning process should be spontaneous and retrieval of encoded information should be implicit and not due to some search for rule conformity (i.e, explicit rule following). It has been suggested that the best way of accessing implicit knowledge may not be by invoking grammaticality judgments. As formal grammaticality judgment instructions may induce a directed rule seeking strategy and therefore push the subjects towards explicit processing. (Manza and Bornstein 1995) A combination of the AGL paradigm with a modified mere exposure paradigm possibly leads to implicit processing. When classification is based on liking (or preference judgments) the participants are not led to reflect on the underlying rule system in the same way as when grammaticality classification is used in the classification phase. If the participants are not informed of the underlying grammar before the classification task is initiated and they still make judgments based on grammaticality, it can reasonably be assumed that the classification performance is influenced by implicit processes.

One of the primary objectives of the behavioural study we conducted was to investigate whether the mere exposure variant of the AGL paradigm can serve as a potentially more appropriate model for language acquisition than the original AGL paradigm. As an additional objective, we wanted to find out what is learned in these types of tasks and possibly how the knowledge is represented. This led to the following experimental questions: Do participants exposed to the AGL strings learn to categorize based on the underlying structure? How do instructions influence performance in AGL tasks? Does repeated testing influence performance? Do factors in the strings besides grammaticality influence the acquisition process?

With these objectives in mind we conducted a behavioural study including 40 participants, which were randomly assigned to one of four groups. Two mere exposure groups were subjected to a mere exposure variant of the AGL paradigm, one was tested in each experimental session and the other was only subjected to the classification task in the last experimental session. This group
was contrasted with a standard AGL paradigm group. A control group was included in the study; they performed acquisition tasks with randomized strings. A typical experimental session was composed of an acquisition phase and a classification phase and each subject participated in one experimental session per day.

Some experiments have shown equal sensitivity for preference judgment tasks and grammaticality tasks, but with “lower quality” knowledge in terms of ability to use the knowledge afterwards (Newell and Bright 2001). We therefore hypothesized that mere exposure groups would not perform as well as the grammatical groups and that it would take longer for them to achieve their maximum score. Another reason for this hypothesis was that the grammar instruction might serve as a motivating factor. We also expected repeated testing to influence performance positively, possibly some helpful strategies might be found in the repeated testing sessions. Without acquisition of grammatical strings it ought to be very difficult to notice the underlying structure of strings, we therefore expected the control group not to learn anything during the course of the test session.
2. Theory

The theory section outlines the paradigm Artificial Grammar Learning in greater detail and alternative theories that try to explain learning and knowledge acquisition in this context. The start point is an outline of some of the fundamental theories concerning memory systems including the division of declarative and non-declarative memory systems.

Memory Systems

Memory refers to the capacity of the nervous system to deal with experience. The most fundamental division of memory is the distinction between behaviour and thought. Memory of behaviour is often called procedural memory, whereas memory of thought is called cognitive memory. Procedural memory is the memory of the things you do, and cognitive memory of the memories expressed by thought. Equivalent names for procedural memory are non-declarative, reflexive or implicit memory. Cognitive memory can be referred to as declarative memory and explicit memory. (Tulving 2000) The terms describing memory processes were originally taken from the information processing framework of the 1960s. In this framework the human brain is characterized as an information–processing device. In this model the mind like the computer receives informational input which it retains for a variable duration and generates outputs in some meaningful form. Acquisition refers to the process of acquiring information and placing it in memory. The acquisition process depends on both external and internal factors. Examples of external factors can be instructions and the character of the to-be-learned material. Internal factors that influence acquisition are for instance attention, motivation, strategies, goals and prior knowledge. (Tulving 2000) Retrieval is the process of recovering previously encoded information. These terms has since widely been used to describe processes involved in describing memory processes. (Brown and Craik 2000)

The distinction between declarative and non-declarative memory is based on converging evidence from studies of experimental animals, neurological patients and normal individuals. In human amnesia memory is impaired against a background of normal intellectual function. The symptom is profound forgetfulness. Amnesic patients cannot form and maintain new long term memories. Amnesia is not a unitary phenomenon and comes in many different forms with different sorts of memory loss. Amnesic patients are often used as subjects for memory research since skills they can acquire are thought to be independent of the declarative memory system, and instead be supported by non
declarative or implicit memory systems. At the same time the recognition memory is severely impaired in the amnesic subjects. Recognition memory refers to the fundamental ability to recognize what is familiar as opposed to what is novel. (Squire, Knowlton 2000)

Declarative memory often refers to facts and events that can be recollected consciously. The establishment of such memories seems to depend on medial temporal structures (hippocampus, the entorhinal cortex, the parahippocampal cortex and the perirhinal cortex. Declarative memory is well suited for storing information about single events. (Squire, Knowlton 2000)

Non-declarative memory is thought to be expressed through performance without any requirement for conscious or explicit memory recollection. It is independent of the medial temporal lobe as well as certain diencephalic structures that support declarative memory. Declarative memory is not itself a brain-system construct. Rather it is an umbrella term that encompasses several different kinds of abilities, with the common feature that they’re not declarative. A collective feature of the non-declarative memory abilities is that amnesic behave normally on these types of tasks. The collection of abilities described by the term non-declarative memories is memory for skills and habits, perceptual priming, simple classic conditioning, emotional learning and non-associative learning. Unlike declarative memories non-declarative memories do not depend on a single brain system. These types of memories depend upon structures like the striatum, neocortex, amygdala, cerebellum, and reflex pathways. (Squire, Knowlton 2000).

Priming refers to the enhanced ability to produce or identify stimuli after previously having been presented with them. Priming effects seems to depend on an existing representation of stimuli which on a later occasion allows the stimuli to be processed more easily (Squire and Knowlton 2000). For example, in category learning, subjects show priming-like effects. After exposure to several exemplars of a category, subjects are able to classify new items according to whether they are members of that category or not. AGL can be said to be an example of category learning where subjects show priming-like effects of stimuli and can categorize in different categories. The difference between category learning and AGL is that in category learning subjects are informed of the categories before commencement of the experiment, in AGL subjects are informed of the categories after the acquisition task.

The two most widely studied paradigms of non-declarative memory is the artificial grammar learning paradigm and the serial reaction time paradigm. The serial reaction time task is an example of sequence learning. In the classical setup of this task subjects respond as rapidly as possible with a key press to a
cue which can appear in any one of four locations on a computer screen. Improved reaction time for sequences and decreased reaction times for randomized sequences indicate that learning has taken place. (Shacter, Curran 2000)

Distinguishing between conscious and un-conscious retrieval is described by the terms explicit versus implicit memory. In this context consciousness refers to the awareness of the relation between the current experience (and activity) and the original learning or acquisition episode. Explicit memory is the clear relation perceived in everyday situations where the individual recollects a previous event and is aware of that the experience in the present situation is influenced by the earlier experience. Explicit memory is contrasted with implicit memory. Implicit memory is defined as retrieval of stored information in the absence of awareness that the current behaviour and the experience have been influenced by previous experience. This distinction applies only to the final stage of a typical memory task, which is during retrieval. This is because there is no difference between explicit and implicit acquisition and there is as yet no known way to distinguish between explicit and implicit storage. (Tulving 2000)
Artificial Grammar Learning

Reber (Reber 1967) suggested that humans are able to learn abstract rules in an implicit fashion while being presented with for example consonant strings. He first coined the term “implicit learning” to describe this ability to learn complex structural information in the absence of awareness. Reber defined implicit learning as the learning process by which the subjects come to respond properly to the statistical regularities in the stimulus sample. It is the process by which an individual develops efficient respond patterns to the structural regularities inherent in the stimuli. In contrast to implicit memory, which primarily emphasized limited awareness and/or intention during retrieval, implicit learning primarily emphasized limited awareness/intention during acquisition. In the original experiment, subjects learned information of the lawfulness of the stimulus sequences in a memorization task. Reber’s explanation was that information was abstracted from the environment without recourse to explicit strategies for responding or systems for recoding the stimuli. Reber’s continued studies (Reber 1989) show that subjects perform above chance level on the AGL test tasks in spite of their lack of reported explicit knowledge of the grammar.

As already noted, the AGL paradigm (Reber 1967, Gordon and Holyoak 1983, Whittlesea and Dorken 1993) include two different phases, an acquisition/training phase and a classification/test phase. During the acquisition phase letter strings derived from a finite state grammar are presented to the subjects. Subjects are asked to keep the strings active in short-term memory and after a short delay write them down on a piece of paper or type them on a computer. In the subsequent test phase, when presented with new strings (i.e., not previously seen), some of which were grammatical and others non-grammatical, subjects are asked to make a judgment whether the strings are grammatical or not. The subjects are instructed to trust their instant feeling about the strings and not to think rationally.

The AGL paradigm was initially used to explore implicit memory processes but has since been used to find the answers to primarily the following questions (where the last one is relevant if you choose to use the paradigm as a model for natural language acquisition):

1. What is the nature of the information acquired during learning, how is it represented, and how is it processed and put to use in for example the classification task?
2. Are implicit or explicit learning mechanisms involved during acquisition?
3. How is classification and transfer performance achieved, implicitly or explicitly, rule-based or based on associative chunk/fragment strength/sequential contingencies?

4. Is AGL a relevant model for first language (L-1) or second language (L-2) learning?

The paradigm has been manipulated in a number of ways in order to find the answers to these questions. Examples of manipulations are experiments with infants and auditory input (Gomez and Gerken 1999). Another example is transfer studies with different surface representations (letter set) between training and test (Reber 1967, 1969). The nature of the representations has been explored using the parameter Associative Chunk Strength (ACS) (Meulemans and van der Linden 1997). The latter is defined as a measure of the congruence similarity of bigrams and trigrams of the acquisition strings compared to the classification strings.

**Formal Languages**

The stimulus material in an AGL experiment is mostly based on a formal language generated from a finite state automaton. One component in the definition of a formal language is its finite lexicon (alphabet) \( V \) of terminal symbols, \( V = \{ t_1, t_2, ..., t_N \} \). The set of all possible finite symbol strings that can be generated from the alphabet \( V \) is given by Kleene-star operator \( V^* = \{ \emptyset, t_1, t_2, ..., t_N, t_1 t_1, t_1 t_2, t_1 t_3, ..., t_k t_k t_k, ... \} \). A formal language \( L \) over \( V \) is then defined as a subset of \( V^* \), \( L \subseteq V^* \); a symbol string \( s = t_k t_k t_k t_k \) is well-formed or grammatical if and only if \( s \in L \). (Lewis and Papadimitriou 1981)

A finite state automaton has three components:

1. A finite set of states, some of which are designated as the initial states, called the start state and some of which are designated as final states.
2. An alphabet of possible input letters.
3. A finite set of transitions that tell for each state and for each letter of the input alphabet which state to go to next.
Learning by liking – a Mere Exposure version of the AGL paradigm

Fig 1: The transition graph representation of the finite state automaton used in the experiment, examples of grammatical strings are (MVRXM, VXVS, VXRRRRRM) examples of non-grammatical strings are (MXSSV, VRXSSS, SVRM).

The automaton works by being presented with an input string of letters that it reads letter by letter starting at the leftmost letter. Beginning at the start state, the letters determine the sequence of the states. The sequence ends when the last input letter has been read. A finite automaton either accepts or rejects a string of letters. A language defined by a finite state automaton is the set of strings that the automaton accepts. (Cohen 1997) The finite state automaton can also be seen as a string generator, where it by going through the automaton in correct manners end up in a finite state, and has generated a string.
Representation Systems

An important objective in the study of AGL is to investigate the characteristics of the acquired knowledge (e.g., symbolic or representations based on distributional information) and how this knowledge is processed and put to use in the classification task. The fact that subjects perform above chance on the classification tasks might (potentially) be explained by anyone of the following suggested hypotheses or theoretical frameworks. (Johnstone, Shanks 2001)

1. Rule-based accounts - symbolic cognitive representations
The rule-based symbolic account suggests that participants acquire (at least some) abstract rules of the grammar during the acquisition phase and are able to use these rules during classification. Thus, it is suggested, that the knowledge is represented in a general abstract form. These representations are assumed to contain little if any information pertaining to specific stimulus features. Evidence for this sort of abstract rule representations comes from transfer studies where subjects are able to categorize according to underlying structure and rules on a different surface representation (e.g., different letter set). (Reber 1967, 1969)

2. Exemplar-based accounts
In the exemplar-based account it is thought that the participants learn to classify test items by representing exemplars, which is, storing the particular strings presented during acquisition in memory. Subsequently, during the classification task subjects base their judgment on a similarity measure. The grammaticality effect is then hypothesized to arise because grammatical test items are more similar to the training stings than non-grammatical items. (Vokey and Brooks 1994)

3. Fragment or chunk (n-gram) accounts
The n-gram account suggests that (often, statistical or distributional) properties related to sequential regularities/contingencies are acquired during the training phase. Subsequently, the subjects base their judgments on similarities in a distributional or relative frequency sense on the bigram and trigram level between training items and test items. For example, in support of this proposal, Perruchet and Pacteau (Perruchet and Pacteau 1990) were able to show that participants trained on bigrams used to construct the grammatical strings were able to classify novel test strings above chance. However subjects trained on bigrams were unable to categorize other violation types, for example those that related to the order of the bigrams and trigrams. In contrast, the participants appeared only to classify items based on detected letter pairs.
4. Dual mechanism accounts
The dual mechanism account postulates that the acquired knowledge is represented both in terms of abstract rules and fragments. This suggestion is thus a combination of representation system 1 and 3. For example, Meulemans and van der Linden (Meulemans and van der Linden 1997) suggested that these two mechanisms are at work at different time points in the acquisition process and may play different roles depending on the particular conditions of an AGL experiment. They suggested that the size of the grammar and the number of strings in the training set may be important experimental parameters that effecting the relative importance of a rule-based and fragment-based representation. In short, the individual performance can be accounted for in terms of both rule abstraction as well as the acquisition of chunk (or fragment) information (i.e., so-called chunking).

5. Episodic processing account
As an alternative account to the one briefly sketched above, Whittlesea and Dorken (Whittlesea and Dorken 1997) suggest that subjects actively engage with the training stimuli to meet the demands of the task instead of just passively processing the acquisition stimuli. They argued that episodic processing is engaged during the acquisition phase in addition to abstract rule learning. Their perspective thus suggests that the learning process is not merely stimulus driven but is dependent of the instruction and the type of acquisition task. Participants can utilize the same acquired knowledge explicitly or implicitly depending on whether they understand the relationship between processing fluency and the knowledge they acquired by processing training items in particular ways.

**Explicit or Implicit Processing**

The importance of the distinction between explicit and implicit knowledge in AGL task has increased and figures prominently in contemporary AGL research. It has been argued that implicit learning leads to an abstract representation whereas explicit learning results in exemplar storage. (Shanks, Johnstone, Staggs 1997) Whittlesea and Wright (Whittlesea and Wright 1997) claim that the subject’s intent is a vital factor in most, if not all, learning tasks. The fact that the demand is unspecified does not make subjects passive but rather forces them to make decisions of how they should deal with the stimuli. Implicit learning is therefore not qualitatively different from explicit learning. They’re both the result of the interaction of the stimuli and processing conducted in compliance with the subject’s current intention. Subjects encode stimuli the way they experience them, and how they perceive them depends on how they process
them. There is only one type of learning and on every occasion of learning some aspects will be explicit and some implicit. In contrast to these suggestions, Knowlton and Squire (Knowlton and Squire 1996) were able to show equally good performance on the artificial grammar tasks, both in its classical and its transfer version, with controls and amnesic patients despite the fact that the recognition memory for letter chunks was severely impaired in amnesic patients. Knowlton and Squire’s results indicate that good performance on classification tasks cannot be accounted for by the explicit knowledge of permissible chunks or any other type of episodic or declarative knowledge, since the amnesic patients lack this form of knowledge. This suggests that the episodic processing account is problematic since the results of Knowlton and Squire provide an important indication that the performance on AGL tasks is independent of the declarative memory system. Currently there is no evidence to link grammar learning directly to particular neural mechanisms beyond the data from amnesic patient’s indication that medial temporal regions are not necessary for such learning to occur. (Squire, Knowlton 2000)
The Mere Exposure Paradigm

The claim that the classification performance in the AGL paradigm is related to an implicit learning process is still an open question and the claim has been criticized by some researchers, as indicated above. It has been argued that the AGL task engages explicit operations and therefore it is unclear whether the task is implicit or not. Manza and Bornstein (Manza and Bornstein 1995) suggest that the best way of accessing implicit knowledge may not be grammaticality classification; as such instructions may increase the likelihood that subjects will engage in explicit processing in addition to implicit processing. Instead they suggest that no explicit reference should be made to the acquisition phase of the AGL paradigm in order to ensure that the classification phase remain implicit with a greater probability. Several ways to ensure the implicit character of the paradigm have utilized indirect measurement strategies. These are indirect in the sense that the subjects are unaware of the underlying grammar and are not instructed to try to acquire it at any point of the AGL experiment. Several studies have reported equal sensitivity to grammatical knowledge when using direct as well as indirect measurement approaches (Buchner 1994).

One indirect measurement method is the mere exposure paradigm. It is based on the finding that repeated exposure to a stimulus causes an increased preference or liking of that stimulus as compared to a novel one. The original finding was reported by Zajonc (Zajonc 1968). Rather than being based on an affective process, the mere exposure effect is believed to be based on a cognitive process where stimuli presented below the threshold for awareness are perceived easier and causes processing fluency. This fluency is interpreted or perceived as a preference or liking bias. (Tulving, Craik 2000) If subjects are unaware of exposure to stimuli they cannot become aware of the relationship between a later memory and a prior acquisition experience. Mere exposure experiments often used geometric shapes that were presented for a sufficiently short exposure time for subjects to report not to have seen or been able to identify the stimuli. At a subsequent test occasion the subjects are significantly more likely to choose the previously presented items compared to similar but new items based on a liking/preference criterion. This paradigm has been used in different experimental settings, and has been shown to yield robust preference results. Thus the effect is robust and can be viewed as an indirect measure of a representation of prior experiences (Whittlesea and Wright 1997). Gordon and Holyoak (Gordon and Holyoak 1983) were the first to adapt the AGL paradigm to a mere exposure version. There were able to show higher affective ratings for grammatical strings compared to non-grammatical strings using a 7 step scale for classifications. This task did not involve memorization of strings but the subjects merely looked at the strings. Their results indicated that the material
had been implicitly processed and thus generalized the observation of the mere exposure effect also in the AGL setting. In AGL experiments it is not possible to present items for such a short time as used in the original mere exposure experiments. Instead, in the mere exposure version of the AGL paradigm, subjects are not informed about the existence of the underlying grammar in the classification phase, and are asked to judge whether they find the string likeable or not. Two possible explanations for the above chance classification performance also in the mere exposure version of the AGL paradigm are related to the suggestion that subjects base their preference judgements on perceptual fluency and ease of processing.

The results reported by Manza and Bornstein (Manza and Bornstein 1995) on the mere exposure version of the AGL paradigm provided initial support for the suggestion that the liking/preference task may be a valid measure of subject’s implicit learning ability. In that study none of the participants in the liking group reported awareness of an underlying rule system or systematic generation principles for the stimuli used in their experiment. Furthermore, Newell and Bright (Newell and Bright 2001) argued that their results indicated that the knowledge acquired with mere exposure instructions might not be as robust as the knowledge acquired in the grammatical judgment version. They were not able to show that the performance of subjects participating in their mere exposure version, and by implication the acquired knowledge, transferred or generalized to a testing situation utilizing a new surface form, that is, when a new string alphabet was used. They were also unable to show improved performance when subjects practiced on bigrams and trigrams in the mere exposure version. Newell and Bright used completion tests after the judgment task and found that subjects with mere exposure instructions were less able to perform well in that type of task. They therefore considered the mere exposure version of the AGL paradigm to be a less sensitive measurement of acquired knowledge compared to the classical AGL paradigm task.

**Associative Chunk Strength**

Reber (Reber 1989) suggested that learning in the AGL paradigm is a consequence of an implicit rule induction mechanism. These rules that remain inaccessible to consciousness are used for grammaticality judgments. This view has however been challenged on a number of issues. It has been argued that participants do not learn any rules at all and that the above chance performance may be explained without reference to an implicit perspective on the acquired knowledge. For example, some researchers have suggested that all knowledge
can be accounted for by knowledge of chunks, or part of strings (e.g., bigrams and trigrams). The proponents of such alternative explanations suggest that (statistical or distributional) knowledge of chunk frequency is enough to account for the individual performance (Perruchet and Pacteau 1990). In order to investigate these issues in greater detail, many scientists in the field have included the so-called Associative Chunk Strength (ACS) as an explicit parameter in their experimental designs. ACS is a statistic measure of the degree to which test strings can be associated with the strings in the training set. In principle it is a measure of the overlap of letter fragments (chunks) of bigrams and trigrams between test and training strings weighted by the number of times a chunk occurred in the training strings. The ACS can be calculated by measuring the chunk frequency for any chunk located in the string (global ACS) and the measure of the repetition of chunks located at the beginning and end of the string (ACS for anchor positions, i.e., initial and final position in a string). It has been shown that the factors ACS and grammaticality can interact and help subjects when classifying strings. Meulemans and van der Linden (Meulemans and van der Linden 1997) were able to show that strings that were both grammatical and high ACS were most often endorsed by subjects. They were also able to show that when subjects were trained on only 16 strings their classification was dependent on ACS but when trained on 125 strings (out of a total of 150 possible strings) their classification was not based on ACS but on grammaticality. This could of course also be an effect of that different features influence the knowledge early and late in the acquisition process. They therefore suggested the Dual Mechanisms Account presented above.

**Control Groups in AGL Experiments**

A control group is defined as a group that is exposed to all the conditions of an investigation except the experimental variable (independent variable). The control group should in all aspects be treated like the experimental group (Solso, Johnson, Beal 1998). In the AGL paradigm there is a need to control for several effects including the testing as such as well as particularities of the stimulus material used in a given experiment. In order to determine that participants learn from acquisition we have to rule out that they can learn from the test situation. There are different approaches in using control groups in AGL, including engaging the controls in merely the classification phases. Alternatively you can engage the control group in the acquisition task but use randomized strings, which thus lack any internal structure. In this way the experiment is similar and matched between groups in relevant respects.
Reber and Perruchet (Reber and Perruchet 2003) showed in an experiment testing controls that endorsements of control participants can be highly systematic. The fewer different letters and the fewer the repetitions a consonant string had the more likely it was to be endorsed. Control participants endorsed letter strings less if it contained repetitions. Other forms of regularity, like repetition of blocks, were more readily endorsed. It was suggested that because the subjects are told about the grammatical structure being difficult to understand the participants are less likely to presume that the grammatical items are composed of obvious or simple feature such as for example simple letter repetition.

If one take the view that learning represents a replacement of one set of biases with another, then it becomes necessary to investigate the initial or prior biases the subjects start out with in any given experiment. (Dienes and Altmann 2003) Starting levels for the subjects have been expected to be chance level, 50 % correctly categorized strings. The reason for this is that subjects without any knowledge of the grammar should perform and random-level, hence being as likely to accept as to reject each string. Recent experimental results have shown that the departure from a theoretical chance level can be quite substantial. For example, Reber and Perruchet reported an initial bias yielding a performance level of 60 % correctly categorized strings. Bias estimation, before any exposure to acquisition strings is a way of coping with this uncertainty.

**AGL and Language Acquisition**

AGL has been proposed as a potentially relevant laboratory model for language acquisition, for first language acquisition in infants as well as for second language acquisition in adults. For example, Gomez and Gerken (Gomez and Gerken 1999) conducted an auditory AGL experiment with infants of 10 months. They used auditory stimuli with syllables and presented them to the infants and used the head-turn-preference procedure in both the acquisition and the classification task where they measured the period of time the infant seemed to be interested by the reading of strings focusing on a light from where the voice was heard. They were able to show that infants learned to distinguish new string from those with endpoint violations as reflected by significantly longer looking times to new grammatical compared to ungrammatical strings. Infants were also able to generalize to another vocabulary (i.e., a new surface form).
The Experiment

Based on the theoretical background outlined above we designed and executed a behavioural study that will described in some detail in the following. The present study differed from several of the previous AGL experiments described above in that subjects did acquisition and test tasks for five consecutive days.

The objectives of the behavioural study were to determine if learning occurs after sequential presentation of strings in the preference group, to investigate how repeated testing influence performance and if performance is dependent on ACS, grammaticality or both. In order to achieve this we manipulated the experimental instructions by giving half of the subjects a grammatical instruction and the other half a mere-exposure instruction. The degree of implicitness was manipulated by the differing instructions as well as the number of classification tasks executed by the mere exposure-instruction groups (i.e., classification tasks every day or only on the first or last day). We also manipulated the acquisition stimuli and gave one group randomly generated strings (i.e., random strings not generated from the artificial grammar) during the acquisition phase. In addition, ACS was manipulated (i.e., high and low ACS) that were presented in a balanced fashion during the classification phase (i.e., during testing). All classification tasks contained an equal number of strings with high and low chunk strength as well as grammatical and non-grammatical strings. Thus, in this part, the experimental design reflects a 2 x 2 factorial design.
3. Materials and Methods

The experiment was executed with 40 subjects that came once a day to do acquisition tasks and classification tasks. The characteristics of the subjects as well as the procedure of generating the stimulus material will be described in some detail below. As the instructions were of vital importance in the study emphasis will be made in explaining them as well as the experimental procedures.

Subjects

Subjects were contacted by e-mail and posted ads. They were asked to get in contact with the experiment leader if they were interested in participating in the experiment. In total, forty-one right handed volunteers (18-40 years) were recruited. They all had a minimum of one year of university education. Participants were pre-screened by means of a standard questionnaire for previous or present neurological or psychiatric disease, drug abuse (including nicotine) and any medication use. One subject was excluded due to failure to comply with the test schedule. 40 participants finished the five-day test series and were included in the subsequent data analysis. All subjects were informed orally and in writing about the schedule and the task in which they participated. All participants gave written informed consent. The participants received a small economic compensation following the completion of the whole experiment. The experimental protocol was pre-reviewed and approved by the local Ethics committee.

Stimulus Material

The stimulus material consisted of consonant strings of a length between 5 and 12 consonants. The experimental strings belonged to one of the following three groups: grammatical strings from the Reber grammar of the letter base (M R S V X) (fig 1) (appendix A), randomized strings from the same letter base and non grammatical strings. Non-grammatical strings had syntactic violations in two positions compared to the grammatical ones. Randomized strings (appendix B) had no underlying structure governing the order of the consonants. In the randomized strings all consonant were equally likely in any specific position. The acquisition set (i.e., ‘acquisition’ strings) consisted of 100 strings that were either grammatical or randomized. The classification task strings consisted of six sets of 40 strings half of which were grammatical and the other half non-grammatical relative the artificial grammar (appendix C). None of the strings in
the classification tasks had previously been seen in the acquisition task by the subjects.

The classification task strings were further classified for Associative Chunk Strength (ACS). High ACS means that there is a high commonality with frequent series of 2 and 3 letter sub-strings (i.e., bigrams, trigrams) as observed in the acquisition set and the given the given classification task string, both for general and anchor (first and last) positions. Low ACS entails a low commonality with the most frequent letter combinations. Thus, the strings in the classification task set were classified into four groups based on ACS and grammaticality (high ACS grammatical, low ACS grammatical, high ACS non-grammatical and low ACS non-grammatical). Each classification task included 10 strings from each category, in total 40 strings for each classification test. Both the acquisition task and the classification task were implemented with the Presentation software (http://nbs.neuro-bs.com). Both tasks were computerized as to allow fully automatic test procedures. Individual responses for each item were recorded during the whole experiment. Each day test scores from the acquisition and classification phases were collected as well as each individual’s subjective rating of performance and task difficulty using VAS-scales (i.e., subjective estimate using a visual analogue scale).

**Experimental Groups**

As noted above, the participants were divided into four groups of ten subjects, which were balanced for gender. The participants were otherwise randomly allocated to one of the groups. The experimental procedures for the different groups are summarized in figure 2.
Learning by liking – a Mere Exposure version of the AGL paradigm

Figure 2: Experimental procedure

Subjects from the different groups were not given the opportunity to communicate with one another during the entire experiment. All groups engaged in the short-term memory task during the acquisition phase each day. Groups A-C processed grammatical strings during the acquisition phase while group D processed random strings. Group A and B performed classification tasks based on a preference instruction. Group B performed classification tasks every day while group A was only did so on day 5. After preference classification on day 5, both group A and B received grammaticality instruction and subsequently performed a grammatical classification. Both group C and D were engaged in grammaticality classification during testing each day. On the first day, before the participants had had any exposure to the experimental material, all groups performed a bias test of the kind corresponding to their later tests in order to estimate possible initial response bias pertaining to the instruction and/or the stimulus material, or any other factor not related to the experiment.
Instructions
All procedural instructions were given to the subjects in writing as well as orally by the experiment leader. A short instruction was also presented on the screen at each experimental occasion. As the experimental manipulation in this experiment pertained to the instruction prior to the exposure to the letter strings great care was taken to ensure that identical and constant instructions were given across and within the different groups.

All groups were tested for the effect of the instruction prior to the first acquisition session. The subjects were informed that to provide background information they were to complete a classification task in which they were presented consonant strings on a computer screen. The preference groups were instructed to make a decision based on the preference criterion while the grammatical groups were instructed to base theirs on the grammaticality criterion.

Before commencement of the acquisition task the preference groups (A and B) were informed that the acquisition task was a short term memory experiment and that they were to focus their attention on the presented strings, keep them in memory for a short delay, and then retype them as accurate as possible on the computer keyboard. At this time subjects of group B were informed of the additional task that would follow the acquisition phase. In this task, they were instructed to complete a preference classification on strings that were similar to the ones encoded. The same instruction was given to group A following the acquisition phase during day 5.

The grammatical groups were informed that they were taking part in an implicit memory experiment. Their first short-term memory task was to look and focus their attention at the strings presented on the computer screen and then retype them on the computer keyboard, as for group A and B. The second task would be to process similar visually presented strings that were or were not generated from the grammar underlying the acquisition strings. Their task was to make a decision based on their immediate gut-feeling rather than any elaborate or explicit strategy for classifying the string as grammatical or non-grammatical.

In the preference groups, following the preference classification on day five, the subjects were informed that the strings they had been exposed to had an underlying grammatical structure. They were then asked to complete a grammaticality classification task based on their initial gut-feeling (i.e., the same instruction as for group C and D).
**Experimental Procedure**

Participants of all groups, except group A, were asked to complete the two tasks every day starting with the acquisition task followed by the classification task. During the acquisition task 100 strings were presented on a computer screen. Each string was presented for 5 seconds and the participants were asked to retype the string correctly afterwards. Immediately after the string was shown a screen with the instruction “retype the string” was shown. No time limit was given the subjects for typing in the string, although they were discouraged from taking too much time on each string. The presentation order of the strings was randomized in the acquisition tasks as well as the classification task. Each acquisition task took about 25-45 minutes. The same 100 strings in the acquisition set were shown to the subjects each day during the acquisition phase.

The six classification tests were balanced across subjects, days and groups. All classification tests were identical apart from the instruction given and the particular strings presented. During a classification session 40 strings were presented on a computer screen, one at a time, each string was presented for 3 seconds. The subject then had 2 seconds to make their classification decision and push the corresponding key, based on the grammaticality criterion (grammatical or not) or by the preference criterion (likeable/pleasant or not). Each classification session was 5 min long.

Following each task subjects filled out a VAS form consisting of 4 VAS-scales. Subjects rated perceived difficulty, degree of accuracy, degree of attentiveness and degree of motivation. In the preference judgment task subjects rated the degree of likeability of the test strings instead of degree of accuracy. Post-experimental interviews took place after the subjects had completed all tasks the last day. The focus of the post-experimental interview was to investigate whether the subjects had used any particular strategy during the acquisition or the classification tasks. Questions of what features the subjects found grammatical or preferable were emphasized (appendix D).

**Data Analysis**

The data was analyzed using multi-way analysis of variance (MANOVA) along with t-tests related to specific comparisons between experimental conditions using the Statistica software package. A $P = 0.05$ significance level was used in all cases. Test scores were based on hit rate and endorsement rate. The four different categories of answers hits, correct rejections, false alarms and misses are explained in fig 3. Hit rate is the sum of all hits (accepted grammatical strings) and correct rejections (rejected non-grammatical strings). Hit rate is beneficial when you want to look at tendencies in learning to categorize according to the underlying grammar. Endorsement rate is the amount of all the
strings that are accepted by subjects, hits and false alarms (accepted non grammatical strings). This way of looking at data is beneficial when you want to see how factors like ACS influence performance independently of grammaticality.

<table>
<thead>
<tr>
<th></th>
<th>Accept:</th>
<th>Reject:</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>G:</strong></td>
<td>Hit</td>
<td>Miss</td>
</tr>
<tr>
<td><strong>NG:</strong></td>
<td>False alarm</td>
<td>Correct rejection</td>
</tr>
</tbody>
</table>

Fig 3: Test scores can be divided into four different categories. The first row represents the measure endorsement rate (hits and false alarms) and the diagonal represents the measure hit rate (hits and correct rejections). It should be noted that when looking at tendencies over time in the hit performance, the false alarm performance is the exact opposite curve.
4. Results

The analysis was made using five repeated measures four-way MANOVA with the between factor group and the within factors time, grammaticality and ACS. The same within factors were used in all MANOVAs but different levels of the repeated measure factor time. Test scores were based on hit rate and endorsement rate. The within factor time had 7 possible levels and corresponds to the number of tests, pre and test 1 took place day 1, test 2 day 2 and so on. Test 6 is the grammatical test of the preference groups after having completed the preference test on day 5.

MANOVA 1 compared all groups on the classification test pre and 5 with hit rate, MANOVA 2 compared learning groups (A, B and C) on the classification test pre and 5 with endorsement rate and MANOVA 3 compared the groups B and C on the classification test 1-5 as well as the pre-test (3A with hit rate and 3B with endorsement rate). A complementary MANOVA (MANOVA 4) analyzed the preference groups vs. grammatical groups on the pre-test (4A with hit rate and 4B with endorsement rate). MANOVA 5 compared test 5 and 6 in the preference groups with endorsement rate (before and after their grammaticality instruction). All collected data are summarized in fig 4 with the hit rate measurement.

![Figure 4: Mean values for all groups on all test occasions.](image-url)
In MANOVA 1 all groups were compared using two levels of the within factor time (pre, 5). Results showed main effects for time (F(1,36) = 48.944, P < 0.001) and group (F (3,36) =17.369, P < 0.001). Significant differences between test occasion pre and 5 are interpreted as learning effects. Contrasting of each group separately comparing test occasion pre and 5 showed learning effects for group A (F(1, 36) = 14.56, P = 0.001), B (F(1,36) = 6.472, P = 0.015), C (F(1,36)= 58.28, P < 0.001) but not for group D. Hence groups A, B and C showed learning effects whereas group D did not (figure 5).

Figure 5: Test scores for all groups prior to acquisition and on test occasion 5 (hit rate). Vertical bars denote confidence intervals for each group on each occasion.

Further contrasting was made to explore what subjects of different groups had learned. Group A had significant differences between test results on the two test occasions in classifying the grammatical strings ( F(1,36) = 9,2726, P = 0.0043), also in classifying the grammatical strings with high ACS F(1,36) = 11,0470, P = 0.0020) and non-grammatical strings with low ACS F(1,36) = 6,4848, P = 0.0153). Group B showed significant learning effects in classification of the grammatical strings F(1,36) = 6,340, P = 0.0163), which seems to primarily consist of improvements in the grammatical strings with high ACS F(1,36) = 6,7425, P = 0.0135). Group C showed learning effects in classification of the grammatical strings F(1,36) = 41,497, P < 0.001), both in the high ACS F(1,36) = 41,469, P < 0.001) and the low ACS strings F(1,36) = 20,668, P < 0.001). Group D did not show improvements in either the grammatical or the non-grammatical strings, but in the high ACS grammatical category (F(1,36) = 4,3152 P = 0.0044). Improvements in classification of strings in the endorsement measurement can be seen in figure 6. Contrasting the learning groups on test 5 showed no significant difference between groups. This lack of significant
difference between the preference groups A and B indicates that repeated testing has little or no influence on the final classification performance. Preference instruction groups seem to learn as well as do grammatical instruction groups.

Figure 6: Endorsement rate divided into ACS and grammaticality for all groups.

MANOVA 2 compared learning groups (A, B and C) with the endorsement measurement on test occasions pre and 5, and showed main effects for time (F(1,27) = 6.49, P = 0.0168) and grammaticality (F(1.27) = 46.16, P < 0.001). Interactions were found between group x ACS (F(2.27) = 7.16 P = 0.00319), grammaticality x ACS (F(1,27) = 57.67, P < 0.001) time x ACS (F(1,27) =14.51, P < 0.001). Three way interactions were found between grammaticality x ACS x group (F(2,27) =6.19, P = 0.0061). There was also a four way interaction (F (2,27) =3.67, P = 0.0389). In this MANOVA we can see that ACS does not influence performance independently.

MANOVA 3A had 6 levels of the within-factor time and compared the two experimental groups B and C with hit rate (mean scores are presented in figure 7). The MANOVA showed no significant effect of group. Main effects were found for time (F(5,90) = 20.19, P < 0.001) and ACS (F(1,18) = 7.170, P = 0.015). Interactions were shown between time x group (F(5,90) = 5.456, P < 0.001), ACS x group (F(1,18) = 5.93, P = 0.02), time x grammaticality (F(5,90) = 3.42, P = 0.007). Three-way interactions were found between grammaticality x ACS x group (F(1,18) = 34.91, P < 0.001) and time x grammaticality x ACS (F(5,90) = 2.45, P = 0.039). Contrasting on each measuring time showed
significant differences only on the first test occasion \( (F(1,18) = 17.76, P = 0.001) \). This difference was further examined in MANOVA 4. Contrasting of the results of test session 3 and 6 showed significant differences in both groups \( (F(1,18) = 8.477, P = 0.009) \), whereas test session 4 and 6 does not. This indicates that the learning curves have levelled out and little or only modest additional learning took place after the third experimental session. Thus, the manipulation of instruction and thereby possibly the subjects’ way of organizing the material does not seem to influence the end proficiency for string classification.

![Graph](image)

Figure 7: Test scores for experimental groups B and C on the six occasions. Vertical bars denote confidence intervals for each group on each occasion.

MANOVA 3B had 6 levels of the within-factor time and compared the two experimental groups B and C with endorsement rate. The MANOVA showed no significant effect of group. Main effects were found for time \( (F (5,90) = 3.4246, P = 0.0070) \) and grammaticality \( (F(1,18) = 48,298, P < 0.001) \). Interactions were found between group x ACS \( (F(1,18) = 34,92, P < 0.001) \) time x grammaticality \( (F(5,90) = 20,199 \ P < 0.001) \), time x ACS \( (F(5,90) = 2,45 \ P = 0.0394) \), grammaticality x ACS \( (F(1,18) = 7,17, P = 0.0154) \). Three way interactions were found between group x grammaticality x ACS \( (F(1,18) = 5,93 \ P = 0.0255) \), group x time x grammaticality \( (F(5,90) = 5,45 \ P < 0.001) \). Again when we look at data throughout the week we can determine that ACS does not seem to influence in any easily determined way but interacts with other factors.

In MANOVA 4A significant differences were found when comparing the preference groups with the grammatical groups on test session 1 with the hit rate measurement, that is, before any exposure to strings (figure 8). The preference
groups categorized 57% correctly based on the preference criteria whereas the grammar groups categorized 40% of the strings correctly based on the grammaticality criteria (F(1,38) = 9.183, P = 0.004). A three way interaction group x grammaticality x ACS (F(1,38) = 7.97, P = 0.007) was found.

Figure 8: Test scores of preference groups (A and B) and grammaticality groups (C and D) on the pre-test.

MANOVA 4B used the endorsement measurement and showed interactions for group x grammaticality (F(1,38) = 30.833, P < 0.001) and group x ACS (F(1,38) = 7.969, P = 0.0075). The differences between the preference groups and the grammaticality groups appears to be that the preference groups have significantly more hits (F(1,38) = 19.053, P < 0.001) which seems to be explained by the differences in answering yes to grammatical strings with low ACS (F(1,38) = 26.149, P < 0.001) for the three other categories of strings there were no significant differences between groups independently.

MANOVA 5 compared preference groups on the two test occasions before and after they were informed of the underlying structure with endorsement rate, it showed main effects for grammaticality (F(1,18) = 115.633 P < 0.001) and ACS (F(1,18) = 20.645, P < 0.001). Interactions for group x ACS (F(1,18) = 11.449, P = 0.0033) and grammaticality x ACS (F(1,18) = 11.911, P = 0.0028), three way interactions were found between group x time x grammaticality (F(1,18) = 5.242 P = 0.0343), and time x grammaticality x ACS(F(1,18) = 10.800, P = 0.0041). Contrasting did not show significant differences between test 5 and 6 for both groups in either the grammatical or the non-grammatical strings. It did however
show significant differences for group A in the non-grammatical strings (F(1,18)=5.65, P=.029).

**Qualitative Data**

Qualitative data was collected in the post-experimental interviews and with the visual analogue scale ratings. The results from the VAS-scale have not been thoroughly investigated at this time.

**Acquisition Phase**

Many different strategies for the acquisition task were reported. With very few exceptions subjects of all groups reported using some sort of blocks of consonants in order to remember the strings, mostly three-consonant-blocks. Some subjects reported reading the blocks or sing them and some reported adding vowels to the consonant strings in order to make them into words. Basing the string on the “VRX”-block and then remembering the letters surrounding them, placing fingers on the first three letters on the keyboard and focusing on the rest was other strategies reported. Most of the participants enjoyed being a part of the experiment and rated high on motivation. Some reported some loss of motivation towards the end of the week and these subjects commonly felt a bigger challenge during the first few days of the experiment. The participants of the random group D were nearly as motivated throughout the week as the rest of the participants. Some did report a decrease in motivation during the last two sessions since they did not feel any improvement and the task was as difficult as on the first day. All subjects were eager to talk about their experiences during the acquisition task were they applied a number of different individual strategies. It was harder for them to describe the strategies used during the classification task and what they based their classification decisions on.

**Classification Phase**

Instructions for the preference groups did not involve any reference to the acquisition task prior to the classification task. One participant in group B reported that it was obvious that the acquisition strings had an underlying structure and that we were interested in finding out if they could learn it. The rest of the participants were unaware of any underlying structure and did not connect the two tasks in that way. The preference groups reported that strings that they preferred, and hence judged as such, were strings that they thought they would have been able to remember. They also reported that preferable strings had a certain rhythm to them when they read the blocks to themselves. Some also mentioned specific features like many consecutive “S” as likeable.
The grammar group C’s answers to what they based their judgments on were quite similar to the ones from the preference groups. They also reported recognition, fluency and rhythm as criteria for grammaticality judgments. Group D gave very vague answers when asked what they based their judgment on. None of the participants could really motivate but referred to “gut feeling” and guessing.

It seems that subjects of group C (6 of 10) were more likely to report having used order and placement of blocks in classifying strings; this was very uncommon in the preference groups. In the preference group B only 1 of 10 reported that the order of the blocks played a role in classification. However, with few exceptions subjects reported having used blocks to categorize the strings. When asked to report the blocks some were able to give as many as 7 while others only reported 2. There were no differences in the amount of reported blocks between the different groups. Most of the participants of group D reported having used blocks of three but they could hardly report any of these.
5. Discussion

The results from the present study demonstrate that subjects have the ability to acquire (some form of) knowledge of the underlying generative structure (i.e., the artificial grammar) as reflected in the presented letter strings. However, the nature and characteristics of the knowledge representations are still unclear. This learning is apparent both in groups that were informed about the presence of an underlying structure and those where no such information was provided.

An important difference between our experiment and most AGL experiments (Reber 1967, Gordon and Holyoak 1983, Whittlesea and Dorken 1993) is the use of consecutive days for encoding/decoding sessions. Obviously the effect is that the subjects will have seen the encoding strings five times and have opportunities to refine their strategies for solving the task. Another difference is that for the groups with grammatical instructions the intent of the study is known from the start, which is not the case in the regular AGL paradigm where subjects are informed after acquisition that there is an underlying grammar governing the order of the letters. This is a potential problem especially in comparing the experiment to other experiments executed in the paradigm. It should however be noted that the subjects repeatedly were discouraged from applying explicit strategies but rely on gut-feeling.

Meulemans and van der Linden (Meulemans and van der Linden 1997) were able to show that different mechanisms governed learning depending on the number of acquisition strings. By their measure, our study is large with 100 strings in the acquisition set. Meulemans and van der Linden suggested that the participants are more likely to develop a kind of abstract rule-like representation of the underlying grammar as a result of implicit learning compared to designs with smaller acquisition sets. With respect to the high/low ACS manipulation, Meulemans and van der Linden did not find that ACS had any influence on classification performance in the case of a relatively large acquisition set. In our experiment ACS did not influence the model as a whole. However it appears to be of some importance interacting with other factors (see the “Learning” section).

Initial Bias

Reber and Perruchet (Reber and Perruchet 2003) pointed out that the chance level (50 %) does not always reflect baseline performance and that participants can start at a significantly different classification level in a pre-test estimation of
the sample specific bias and baseline performance. In our experiment we observed differences between preference groups and grammaticality groups at the first baseline performance estimation, with the preference groups classifying above chance level while the grammatical groups classified below chance. The fact that grammatical groups reject grammatical strings to a higher degree than chance indicates that they are sensitive to a different structure than the one represented by the Reber grammar. It could possibly be explained by salient features that belong to the grammar like letter repetition and block repetitions are rejected. Features of this kind might not be part of their mental model for what a grammar is. Instead, the preconceived idea of a grammar might relate to a rather more complex idea of a grammar. Since differences between grammatical and non-grammatical strings are subtle, differing only in two letter positions it is a rather surprising finding. However, little is known about the reasons and factors that influence the departure of the initial classification bias from the prior chance level.

**Learning**

The design of the present study allows us to investigate the learning or acquisition rate of the subjects in the different instruction groups. It also gives some indication of what is learned and how long it takes for subjects to reach the point at which increased classification performance starts to level out. In the present study experimental groups reached this point approximately after three days. This is earlier than we had expected. When looking at the learning groups A, B and C you find that their learning significantly is established by their learning to endorse grammatical strings more often. (figure 6)

The parameter Associative Chunk Strength does not appear to influence performance in the classification tasks independently as indicated by ACS not being a main factor in the endorsement rate MANOVAs (with the exception of MANOVA 5). However ACS seems to interact with the other factors time, group and grammaticality (in MANOVA 2 and 3B). All groups endorse significantly more grammatical strings with high ACS after five days of acquisition. The fact that even group D improves in classifying this category of strings indicates that these strings are the easiest ones to detect. According to Meulemans and van der Linden (1997) the grammatical strings with high ACS are the ones most likely to be endorsed. It should also be noted that the fact that ACS is a factor in the endorsement rate MANOVA 5, is most likely a consequence of the dramatic change in endorsing strings of this category, while other changes are quite modest.

Additional significant differences are found when contrasting scores on the pre test with test 5 for group C in the low ACS grammatical category. Group A shows tendencies of more often endorsing these strings, whereas group B that
start out with a strong bias of endorsing the G-L strings don’t show any tendencies of improving. One possible interpretation for performance in group B is that the local features of these strings is endorsed or rejected at the first test occasion and if they are easy to recognize they may influence the classification decisions in the following test sessions based on the initial classification decision. This in conjunction with the fact that the participants never receive any performance feedback may make it difficult to ‘unlearn’ this response bias introduced in the initial test.

Group B is the only group that shows significant differences in rejecting in the non-grammatical strings with low ACS more often. But tendencies are apparent in as well group A as group C. This category of strings is the one that mostly differ from the other groups since they are neither grammatical nor highly associated with acquisition strings. There is a tendency of differences between the two test occasions in group C in the non-grammatical high chunk strength strings, which is not apparent in any of the other groups. This category of strings thus seems to be the hardest ones to recognize and reject. The reason for this is that these strings because of their high ACS are very similar to the ones that subjects have seen in the acquisition phase. This is probably the reason for the lack of improvements in rejecting the non-grammatical strings with high ACS.

**The Mere Exposure Paradigm**

We were able to show equal sensitivity and classification performance with respect to grammaticality in the classification task in both the mere exposure instruction groups and the grammaticality instruction group. These results are in accordance with Buchner’s study (Buchner 1994), which showed similar results with grammaticality and mere exposure instructions. In Manza and Bornsteins’ study (Manza and Bornstein 1995), also using a mere exposure instruction, none of the subjects reported having noticed an underlying structure. This was also the case in our study. With the exception of one subject in the preference group B, none reported knowing or suspecting an underlying structure or system of rules for generating the strings despite being tested on several occasions. In mere exposure version of AGL experiments no mentioning of the grammar can be made prior to the test task. This group did not know what the underlying rational or purpose of the experiment was. Most of the subjects considered the short-term memory acquisition task to be the main task and did not think too much about the classification task. The preference criterion seems to have discouraged from an elaboration of an explicit strategy as well as being less prone to induce an analyzing strategy.

Newell and Bright (Newell and Bright 2001) argued and reported data indicating that the knowledge subjects base their preference judgments on are less robust
than grammaticality judgments and suggested that mere exposure tasks led to what they called “lower quality” knowledge. They were unable to show transfer effects with mere exposure instructions. Their results also indicated that subjects with mere exposure instructions who in their acquisition phase encoded bi- and trigrams did not perform as well as subjects with grammaticality instructions. In our particular case it is difficult to draw any conclusions along these lines since we did not use any particular method to ensure or measure the quality of the acquired knowledge representations. Manza and Bornstein indicated that the preference based classification led to less pronounced explicit rule knowledge of the grammar compared to grammaticality based version. Our post-experimental interviews gave some support for their finding. Group A and B appeared to be less able to report rule-like strategies used during classification, in particular with respect to the order and placement of the blocks or chunks in a string.

The preference groups seem to learn to classify strings at the same performance level as the grammatical groups. It appears also that they to some extent are less able to report “rules” that governed their decision, and seem to have reflected less on the structure. Knowledge acquired with the mere exposure instructions might therefore reflect or be related to a more automatic mode of processing with less interference or support from explicit rule elaborations. Thus it seems that this particular version of our experiment comes closer to or represents a more valid model of implicit natural language acquisition.

**Representation Systems**

Traditionally memory processes are broadly divided into encoding, storage and retrieval. According to Tulving (Tulving 2000) the distinction between explicit and implicit memory is only relevant for retrieval. Reber on the other hand emphasizes the implicit acquisition process. (Reber 1967) A main focus of the AGL research is the nature of what information is acquired and how this is represented. This research gives some suggestions and several possibilities have been outlined in some detail. Five systems reported to explain performance is abstract rule based representations, exemplar-based representations, fragment accounts, dual mechanisms accounts, episodic processing.

With very few exceptions subjects reported using chunks of letters in the acquisition task as well as during classification. It should be pointed out that if they solely base their judgment on single blocks it would lead to chance performance, since the non grammatical strings also contain the blocks reported by subjects. We are therefore led to believe that subjects have acquired knowledge of a nature that they cannot explain, hence implicit knowledge. It also seems relevant to point out that what subjects report may be the subject’s way to try to conceptualize their experience but it possibly has little to do with
how they actually perform the task. Although it was not empirically tested it seems likely that the few rules reported by the subjects are not enough to account for improvement in performance during the week. If it is indeed the case that most of the processing in this paradigm is implicit it would appear to the subject like they are guessing, but their classification performance is based on acquired knowledge.

It is reasonably clear that ACS is a factor that influences the classification performance. Perruchet and Pacteau (Perruchet and Pacteau 1990) suggest that the subject performance, and by implication the knowledge acquired in AGL tasks, can be accounted for by reference to knowledge of frequency of occurrence of fragments of letters. The most convincing evidence for implicit processing in this paradigm comes from Knowlton and Squire who were able to show equally good performance with amnesic patients as with controls. Their results suggest that whether subject achieve fragment knowledge or rule knowledge it is implicit. Thus we are led to believe that fragmentary knowledge of some kind, implicit or explicit, or both are present. If there is also additional representation systems, like rule based or exemplar based, present is not determined.
6. Conclusions

In this behavioural study we wanted to investigate whether the mere exposure variant of the AGL paradigm is a better model to describe natural language acquisition than the original AGL paradigm. We were able to show that subjects learn to classify strings at the same level of performance independently of whether the classification criterion was based on preference or grammaticality. In other words, the correlation between preference/non-preference and acceptance/rejection and the actual grammaticality status of the items were similar in the two instruction groups. It has been suggested that the task of learning an artificial grammar is a potentially relevant laboratory model for investigating aspects of language learning in infants (Gomez & Gerken, 2000), exploring key differences between human and animal learning relevant to the narrow faculty of language (Hauser et al., 2002), as well as second language learning in adults (Friederici et al., 2002).

The process of retrieving knowledge put to use in classifying strings seems to be a less analytic task for the preference groups compared to the grammatical groups, at least as subjectively reported. This provides some support for the suggestion that the mere exposure version of the AGL paradigm is a more appropriate model for natural language acquisition. In addition, we wanted to investigate what is learned and possibly the mode of learning in the different versions of the AGL paradigm used in this study. It seems that the subject performance is not easily explained by one representation system. Subjects base their classification on knowledge of chunks, at least if the subjective reporting is to be taken at face value. Since explicit knowledge of chunks is not enough to account for performance we conclude that there is some sort of implicit retrieval involved. Whether the retrieved knowledge was of fragmentary or abstract rule kind was not determined.
References


### Appendix A: Grammatical acquisition strings

<table>
<thead>
<tr>
<th>MSSS(\text{VRXRRRRM})</th>
<th>MSSSS(\text{SVRXM})</th>
<th>(\text{VXVRXSVRXV})</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\text{MVRXVRXRRRRM})</td>
<td>(\text{VXVRXSSSSSV})</td>
<td>(\text{MSSSSVRXVRXR})</td>
</tr>
<tr>
<td>(\text{VXSVRXRR})</td>
<td>(\text{MSVRXSSSVS})</td>
<td>(\text{VXSSSVRXRVX})</td>
</tr>
<tr>
<td>(\text{MVRXRM})</td>
<td>(\text{MSVRXSSSVS})</td>
<td>(\text{VXSSSVRXVX})</td>
</tr>
<tr>
<td>(\text{MVRXVRXRRRRM})</td>
<td>(\text{VXVRXSSSV})</td>
<td>(\text{MSSSSVRXVRXV})</td>
</tr>
<tr>
<td>(\text{VXSVRXRR})</td>
<td>(\text{MSVRXSSSVS})</td>
<td>(\text{MVRXSVRXRVX})</td>
</tr>
<tr>
<td>(\text{MSSSSSSSVS})</td>
<td>(\text{VXSSSSSVSX})</td>
<td>(\text{MVRXSVRXRVX})</td>
</tr>
<tr>
<td>(\text{MVRXSVXRRRR})</td>
<td>(\text{MSVRXSSSVS})</td>
<td>(\text{VXSSSSSVSX})</td>
</tr>
<tr>
<td>(\text{MSSSSSVSRXRR})</td>
<td>(\text{VXSSSSSVSX})</td>
<td>(\text{MSSSSSVRXRVX})</td>
</tr>
<tr>
<td>(\text{MSSSSSVRXRR})</td>
<td>(\text{VXSSSSSVSX})</td>
<td>(\text{MSSSSSVRXSV})</td>
</tr>
<tr>
<td>(\text{MSSSSSVRXRR})</td>
<td>(\text{VXSSSSSVSX})</td>
<td>(\text{MSSSSSVRXVRX})</td>
</tr>
<tr>
<td>(\text{VXSVRXRR})</td>
<td>(\text{MSVRXSVRX})</td>
<td>(\text{VXSSSSSVSX})</td>
</tr>
<tr>
<td>(\text{MSSSSSVRXRR})</td>
<td>(\text{VXSSSSSVSX})</td>
<td>(\text{MSSSSSVRXSV})</td>
</tr>
<tr>
<td>(\text{VXSVRXRR})</td>
<td>(\text{MSVRXSVRX})</td>
<td>(\text{VXSSSSSVSX})</td>
</tr>
<tr>
<td>(\text{MSSSSSVRXRR})</td>
<td>(\text{VXSSSSSVSX})</td>
<td>(\text{MSSSSSVRXSV})</td>
</tr>
<tr>
<td>(\text{VXSVRXSSV})</td>
<td>(\text{VXSSSSSVSX})</td>
<td>(\text{MSSSSSVRXSV})</td>
</tr>
<tr>
<td>(\text{MSSSSSVRXRR})</td>
<td>(\text{VXSSSSSVSX})</td>
<td>(\text{MSSSSSVRXSV})</td>
</tr>
<tr>
<td>(\text{VXSVRXRR})</td>
<td>(\text{VXSSSSSVSX})</td>
<td>(\text{MSSSSSVRXSV})</td>
</tr>
<tr>
<td>(\text{VXSVRXRR})</td>
<td>(\text{VXSSSSSVSX})</td>
<td>(\text{MSSSSSVRXSV})</td>
</tr>
<tr>
<td>(\text{VXSVRXRR})</td>
<td>(\text{VXSSSSSVSX})</td>
<td>(\text{MSSSSSVRXSV})</td>
</tr>
<tr>
<td>(\text{VXSVRXRR})</td>
<td>(\text{VXSSSSSVSX})</td>
<td>(\text{MSSSSSVRXSV})</td>
</tr>
<tr>
<td>(\text{VXSVRXRR})</td>
<td>(\text{VXSSSSSVSX})</td>
<td>(\text{MSSSSSVRXSV})</td>
</tr>
<tr>
<td>(\text{VXSVRXRR})</td>
<td>(\text{VXSSSSSVSX})</td>
<td>(\text{MSSSSSVRXSV})</td>
</tr>
<tr>
<td>(\text{VXSVRXRR})</td>
<td>(\text{VXSSSSSVSX})</td>
<td>(\text{MSSSSSVRXSV})</td>
</tr>
<tr>
<td>(\text{VXSVRXRR})</td>
<td>(\text{VXSSSSSVSX})</td>
<td>(\text{MSSSSSVRXSV})</td>
</tr>
<tr>
<td>(\text{VXSVRXRR})</td>
<td>(\text{VXSSSSSVSX})</td>
<td>(\text{MSSSSSVRXSV})</td>
</tr>
<tr>
<td>(\text{VXSVRXRR})</td>
<td>(\text{VXSSSSSVSX})</td>
<td>(\text{MSSSSSVRXSV})</td>
</tr>
<tr>
<td>(\text{VXSVRXRR})</td>
<td>(\text{VXSSSSSVSX})</td>
<td>(\text{MSSSSSVRXSV})</td>
</tr>
<tr>
<td>(\text{VXSVRXRR})</td>
<td>(\text{VXSSSSSVSX})</td>
<td>(\text{MSSSSSVRXSV})</td>
</tr>
<tr>
<td>(\text{VXSVRXRR})</td>
<td>(\text{VXSSSSSVSX})</td>
<td>(\text{MSSSSSVRXSV})</td>
</tr>
<tr>
<td>(\text{VXSVRXRR})</td>
<td>(\text{VXSSSSSVSX})</td>
<td>(\text{MSSSSSVRXSV})</td>
</tr>
<tr>
<td>(\text{VXSVRXRR})</td>
<td>(\text{VXSSSSSVSX})</td>
<td>(\text{MSSSSSVRXSV})</td>
</tr>
<tr>
<td>(\text{VXSVRXRR})</td>
<td>(\text{VXSSSSSVSX})</td>
<td>(\text{MSSSSSVRXSV})</td>
</tr>
<tr>
<td>(\text{VXSVRXRR})</td>
<td>(\text{VXSSSSSVSX})</td>
<td>(\text{MSSSSSVRXSV})</td>
</tr>
<tr>
<td>(\text{VXSVRXRR})</td>
<td>(\text{VXSSSSSVSX})</td>
<td>(\text{MSSSSSVRXSV})</td>
</tr>
<tr>
<td>(\text{VXSVRXRR})</td>
<td>(\text{VXSSSSSVSX})</td>
<td>(\text{MSSSSSVRXSV})</td>
</tr>
</tbody>
</table>
Appendix B: Randomized acquisition strings

RRSXX  MMSRXRRXVX  XRMRRSRVSVXVR
RMSXV  MRRSSRSMRMV  RRSVXRVRXRM
SMMXR  MXXRMRMMXMX  XRSSXRRVXSR
RVMXVR  XXRMSXMMRV  RXSRVXVMXRS
XSSRXV  MVRSSMMRMRR  XMXSRSSVSRVX
MMRMRR  SVMXSSMRMR  SVXMMXRSXSV
XSRRSV  XMMXVMMRMR  XSSMMSMMMXXS
VSMXRMV  MXXMRRSMVR  SXRXVVRMRXRM
XRMRSVX  VRSRXRRMMV  VSRXVRVRSVXV
SXVSMXMXV  XSSRSSRMR  MXXMRRSMXMX
XVRRMSSR  MRRMRXMXSSS  XMSVRMVRMSRS
RRXMXVVV  MRRMRVVRMR  RRMMRMMSXMS
MXRXRRSX  RSRMRXSVXV  XSSXSSMMXVX
CSVSMXMX  VRVVRXRVXV  RXXMRSMSVXX
VVIXMRXXV  RVSSRRVTVMM  SXMSXSMMMMSR
VXXMRSXRM  MRSSRMXMRX  SMXRMXRMRSV
RRMSXRSVR  VSMRSMVRXRV  XXVRRVXMMXV
XMMXSSMS  MMSRXRMXSV  VMVXMSXRXRM
RVXXSVXS  MMVXSSXXVX  RRVXSSRSMX
SSXXSRXV  MVRSSRRXXRS  MMSXRVSVVVR
MVSMXMSR  RVSSRVRSSRX  VRXXVXSMSVS
XRMVXVXXM  RRVRXMRRMVX  VMVSVRVMXRR
VXXMMMSRX  RXRMRSMVRS  RXXSSRSSXRX
MXXSSXSS  XRXXSMMVXXS  XSSXRRSSXXM
RRMMRVMVS  SSVRXRXMMMS  RRXXRMMSXSR
XVXXSVSSS  SVSSXSXXMV  XRMSXXSSXXS
XXSMVMSMS  SMVXSRXMMV  SMVXMSXMRV
MSMVXVRX  VVMSXSMXSVX  MMVXVMMXSM
XRSRXSMMS  RSMXVMXVSSS  VVXXSMXMVXV
SMVMRSXX  VRXRRXSRXRM  VMVRMXXRSSV
VSSXMSRV  RSVXSRMSSM  RSRMVRMRVX
RRRXSSXMMX  SVXVMSRSMRM  MRRXSRVXRMX
MSVXMRSSX  MRRXMSXMMV  MRXMSXMRV
RVMXVMSXSV  VRSSXMMRVSXSS
Appendix C: test strings with classification, test 1 of 6

{"highGt","VXVRXVVRXVRXV"}
{"highGt","VXSSVRXVS"}
{"highGt","VXVRXSSSVRXR"}
{"highGt","VXSSVRXSSSV"}
{"highGt","VXVRXSSSVRXR"}
{"highGt","MSSSVRXVS"}
{"highGt","VXSVRXXSSVS"}
{"highGt","VXSSVRXVRXV"}
{"highGt","MSSSSSSSSSV"}
{"highGt","VXSSVRXSSSV"}
{"highGt","VXSSVRXSSSV"}
{"highGt","MVRXSVRXSSSV"}
{"highGt","VRSXSVRXSS"}
{"highNGt","MMSSVRXSS"}
{"highNGt","VXRXSSSVRXR"}
{"highNGt","VXRXSVRXRVX"}
{"highNGt","VRMSVRXRVXV"}
{"highNGt","MRXSSSSSSSV"}
{"highNGt","MVRXSVRXSVR"}
{"highNGt","VMSSVRXSS"}
{"highNGt","VXSSVRXSVV"}
{"highNGt","VXSVRXXSVRS"}
{"lowGt","VXVRXRRR"}
{"lowGt","VXVRXRRRR"}
{"lowGt","VXVRXRRRRRM"}
{"lowGt","MVRXRRRRRR"}
{"lowGt","VXSVRXXRRRRR"}
{"lowGt","MSVRXRRRRRRR"}
{"lowGt","VXVRXRRRR"}
{"lowGt","VXRRRRRRM"}
{"lowGt","VXRRRRRRRRRR"}
{"lowGt","VXRRRM"}
{"lowNGt","VXRRRSRXM"}
{"lowNGt","VXVRXRRSRXM"}
{"lowNGt","VXSVRXSSRRRXR"}
{"lowNGt","VXRXMRRRRR"}
{"lowNGt","MVRXRXMRRR"}
{"lowNGt","MSVRXRXMRRRRR"}
{"lowNGt","VSVRXRRRMR"}
{"lowNGt","MVRXRXSXR"}
{"lowNGt","VXRXMRRRRRRR"}
{"lowNGt","VXRSMV"}
{"lowNGt","VXSVRXXRSM"}
Appendix D: Post interview form Group A

Namn:...........................................

Postintervju AGL - efter fullbordad försöksserie (A)

1. Var det något som påverkade prestationen under försöken (tristess, trötthet)?
   När i så fall? I början, mitten slutet?  Ja  Nej

2. Har din motivation förändrats under försöks seriens gång?  Ja  Nej
   Hur då?

3. Kan du erinra dig något speciellt hos konsonantsekvenserna som du upplevt
   som behagligt eller inte, som du använt i din bedömning av strängarna?
   Ja  Nej
   I sådana fall vad?

   Använde du dig av kombinationer av bokstäver?  Ja  Nej
   I sådana fall vilka?

   Mönster? Vilka?  Ja  Nej

   Placering? Vilka?  Ja  Nej

   Använde du dig av någon speciell strategi vid bedömningarna?  Ja  Nej

6. Kan du erinra dig något speciellt hos konsonantsekvenserna som du ansåg
   vara grammatisk eller inte under den grammatiska bedömningen?
   Ja  Nej
   I sådana fall vad?
Användade du dig av kombinationer av bokstäver?

Ja        Nej

I sådana fall vilka?

Mönster? Vilka?

Ja        Nej

Placering? Vilka?

Ja        Nej

Användade du dig av någon speciell strategi vid bedömningarna?

Ja        Nej

7. Skilde sig preferensbedömningarna och grammatikalitetsbedömningarna mycket åt för dig? I sådana fall hur?

Ja        Nej
Learning by Liking - a Mere Exposure version of the AGL paradigm

Abstract
The artificial grammar learning (AGL) paradigm has been intensively researched since the 60-s. In general, these investigations attempt to study the implicit acquisition of structural regularities. Among other things, it has been suggested that the AGL paradigm can serve as a model for the process of acquiring a natural language. Thus it can serve as a well-controlled laboratory task that might be used to understand certain aspects of the process of language acquisition. For example the AGL paradigm has been used in an attempt to isolate the acquisition of syntactic aspects of language. Several experimental studies show that the participants acquire knowledge of the underlying rule system since they are able to differentiate grammatical strings from non-grammatical ones. It has been argued that the traditionally conducted AGL paradigm with grammaticality instructions might make the task explicit, at least during the test phase. In order to imitate the language learning process as close as possible, to rule out the possibility of an explicit component during the testing phase (i.e., keeping the retrieval process implicit) and to rule out explicit rule conformity or rule following, we modified the classical AGL paradigm. In a behavioural study we combined the AGL paradigm with an altered mere exposure paradigm in an attempt to better model aspects of language acquisition. We were able to show that subjects, classifying under mere exposure instructions, categorize grammatical and non-grammatical strings just as well as those solving the classification task with the grammaticality instructions. This indicates that the mere exposure version might serve as a more appropriate model for language acquisition.

Keyword
artificial grammar learning, AGL, mere exposure, implicit memory, language acquisition